Statistical mechanics of attractor neural networks and self-consistent signal-to-noise analysis:

analog neural networks with nonmonotonic transfer functions and enhancement of the storage capacity

Department of Applied Physics, Tokyo Institute of Technology
Masatoshi Shiino
Department of Electronics, Tokai University
Tomoki Fukai

The self-consistent signal-to-noise analysis (SCSNA) we have recently deveoped is a systematic method to deal with the equilibrium properties of analog neural networks of associative memory with a general type of transfer functions. The method is based on a self-consistent treatment of the local fields of neurons to extract the self-interaction part and gaussian noise by means of a renormalization procedure. Applying the SCSNA to analog networks with nonmonotonic transfer functions in which the updating rule is given by continuous time dynamics and the synaptic couplings are formed by the Hebb learning rule with unbiased random patterns, we have found the networks to exhibit remarkable properties leading to an improvement of network performances under the local learning rule ; enhancement of the storage capacity and occurrence of errorless memory retrieval with an extensive number of memory patterns . The latter is due to the vanishing of noise in the local fields of neurons which is caused by the functioning of the self-interaction part of the local field in combination with sufficiently steep negative slopes in the transfer functions.

1. Introduction

Neural networks are nonlinear dynamical systems of interconnected neurons, which perform various types of parallel information processing such as learning, memory retrieval, pattern classification, thinking, and so on. All of those functions of neural systems should be based on some cooperative behaviors exhibited by a large number of neurons. Associative recall of memory is one of the typical cooperative phenomena of nonlinear systems, such that certain attractors called retrieval states are formed by a network which stores memory patterns in the synaptic connections by means of a certain learning rule and memory retrieval is represented by the dynamics settling into the retrieval state close to a given initial pattern. The dynamics governing the time evolution of a network can be stochastic or deterministic, depending on whether a network is subject to noise or not.

In the case of stochastic dynamics, attractors of a network can emerge giving birth to the so-called ergodic components only when the number of neurons N goes to infinity (thermodynamic limit). Then an appearance of the ergodic components is called the phase transition. When a stochastic network system satisfies the condition of detailed balance with symmetric interactions between neurons as in the case of the Glauber dynamics with the usual Hebb learning rule (Hopfield model), one can employ statistical mechanical approach to study the properties of the ergodic components corresponding to the retrieval states, owing to the existence of an energy function. In fact, Amit et. al.[1] have succeeded in analyzing the dependence of the storage capacity (i.e. upper bound of the number of stored memory patterns per neuron) on the stochastic noise or temperature for the Hopfield model of the Ising spin (binary neuron) neural nerworks within the framework of equilibrium phase transitions in the spin glass theory[2].

The availability of the statistical mechanical theory is not restricted to the case of the Ising spin networks but applies to a certain class of deterministic analog neural networks[3], which are characterized by the transfer functions representing input-output relations of graded-response neurons. A neural network with a monotonically increasing transfer function updated by continuous-time dynamics has an energy function in the case of symmetric synaptic interactions. The existence of an energy function allows one to analyze the storage capacity of the network using the replica symmetric theory in spin glasses[4,5]. An analog network with a hyperbolic-tangent transfer function, which exhibits roughly the same size of storage capacity[4] as obtained by Amit et.al.[1] , has been found to be closely related with the Ising spin networks with the only difference being the presence of the Onsager reaction term [6]in the TAP equation of the latter networks[4,7,8,9].

When, on the other hand, a network has a transfer function which is not monotonically increasing or has asymmetric synaptic connections, it has no longer an energy function and defies use of statistical mechanical approach to investigating the network properties even when it works as an associative memory. Nonmonotonic transfer functions may appear when one considers effective transfer functions that describe overall input-output relations of possible local clusters of neurons serving as a functional unit in the information processing of phisiological nervous systems.

We have developed a systematic method of analyzing the storage capacity of analog networks with a general type of transfer functions to cope with the difficulty of studying networks without energy functions[10]. The new method of the self-consistent signal -to -noise analysis (SCSNA) has its basis in the self-consistent splitting of the local fields of neurons into signal and noise parts by means of a renormalization procedure and yields a set of order-parameter equations describing the retrieval states of the network[10,11]. An essential feature of the SCSNA is the extraction of the effective self interaction part leading to an output proportional term in the local fields. It is noted that if applied to the case of monotonic transfer functions, the SCSNA gives the same result as obtained by replica calculations[4].

Applying the SCSNA to continuous time analog neural networks with a certain type of nonmonotonic transfer functions[11,12,13,14,15], we have found, in addition to an appreciable enhancement of the storage capacity relative to the commonly known value 0.138, quite interesting and remarkable phenomena to occur due to the presence of output proportional term in the local fields. Nonmonotonic transfer functions can exhibit an onset of errorless memory retrieval under the local learning rule of the Hebb type as a result of a phase transition with respect to the width of the local field distribution. It has been so far believed that networks which store an extensive number of patterns through the local learning rule can not be free from a finite fraction of errors in its retrieval states, which increase as storage ratio is increased [1].

2. Self-consistent signal-to-noise analysis of analog neural networks

The dynamics of the network of N analog neurons with a transfer function F is assumed to be given by

$$\frac{d}{dt}u_i = -u_i + \sum_{j=1}^{N} J_{ij}\, F(u_j) \ , \qquad i = 1 \ldots N \ , \tag{1}$$

with u_i representing membrane potential of neuron i and J_{ij} synaptic connection from neuron j to i with the Hebb learning rule

$$J_{ij} = \frac{1}{N} \sum_{\mu=1}^{p} \xi_i^{(\mu)} \xi_j^{(\mu)} \quad (i \neq j) \quad , \quad J_{ii} = 0 \tag{2}$$

where $\left\{\xi_i^{(\mu)}\right\}$ ($\mu = 1 \ldots p$, $i = 1 \ldots N$) denote p sets of uncorrelated random patterns.

Assuming the existence of equilibrium solutions to (1), we are concerned with solving a set of equations $x_i = F(\sum_j J_{ij} x_j)$, $i = 1 \ldots N$, for output of neurons ($x_i \equiv F(u_i)$) to obtain the retrieval solution having a significant overlap with a particular pattern, say pattern 1, of the form $m^{(1)} = O(1)$, $m^{(\mu)} = O(1/\sqrt{N})$ for $\mu \geq 2$.

Here $m^{(\mu)}$ denote the order-parameter overlaps

$$m^{(\mu)} = \frac{1}{N} \sum_j \xi_j^{(\mu)} F(u_j) \ , \quad \mu = 1 \cdots p \tag{3}$$

We define another type of overlaps $g^{(\mu)}$, the tolerance overlaps, measuring the degree of the pattern retrieval more appropriately than the order-parameter overlaps in the case of transfer functions which differ much from the sigmoidal ones (11):

$$g^{(\mu)} = \frac{1}{N}\sum_i \xi_i^{(\mu)} \,\mathrm{sgn}\, u_i \ . \tag{4}$$

A perfect recall of memory of pattern 1 should be specified by the condition that $\mathrm{sgn}\, u_i = \xi_i^{(1)}$ for every i, implying $g^{(1)} = 1$.

The SCSNA extracts pure noise part in the local field h_i of neuron i , which is written as

$$h_i \equiv \sum_j J_{ij}x_j = \xi_i^{(1)} m^{(1)} + \xi_i^{(\mu)} m^{(\mu)} + \sum_{\nu \neq \mu,1} \xi_i^{(\nu)} m^{(\nu)} - \alpha x_i \ , \quad \mu \neq 1 \tag{5}$$

, by means of a renormalization procedure involving the self-consistent decomposition [10,11]

$$\sum_{\nu \neq \mu,1} \xi_i^{(\nu)} m^{(\nu)} = \overline{z_{i\mu}} + \gamma x_i \ , \qquad \mu \geq 2 \ , \tag{6}$$

where we have given μ-term special treatment for a technical reason and $\overline{z_{i\mu}}$ represents pure noise. It is noted that the so-called conventional treatment implies $\gamma = \alpha$, giving rise to 0 mean of the naive noise $\sum_{\nu\geq 2}\sum_{j\neq i}\frac{1}{N}\xi_i^{(\nu)}\xi_j^{(\nu)}x_j$.

The result of the renomalization procedure which computes $m^{(\mu)}$ based on (5) to substitute into (6) is that γ is given as $\gamma = \frac{\alpha}{1-U}$, and the noise $\overline{z_{i\mu}}$ is almost independent of μ and i, obeying an identical gaussian distribution with mean 0 and variance σ^2. Here U and σ are given by $U = \langle\langle \frac{\partial x_i}{\partial \overline{z_{i\mu}}} \rangle\rangle$ and $\sigma^2 = \frac{\alpha}{(1-U)^2}\left\langle\!\left\langle x_i^2 \right\rangle\!\right\rangle$ with x_i representing the renormalized output and $\langle\!\langle\ \rangle\!\rangle$ denoting average over $\xi_i^{(1)}$ and gaussian noise $\overline{z_{i\mu}}$. These two equations together with the one for m constitute the SCSNA order-parameter equations which self-consistently determine m, U, and σ^2 to describe the retrieval states. With setting $\sigma^2 = \alpha r$ and $z = \overline{z}/\sigma$ they read [10,11]

$$m = \int_{-\infty}^{\infty} dz \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) Y(z), \quad \left(1-U\right)^2 r = \int_{-\infty}^{\infty} dz \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) Y(z)^2, \quad U\sqrt{\alpha r} = \int_{-\infty}^{\infty} dz \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) zY(z), \tag{7}$$

in the case of an odd transfer function, where the renormalized output Y(z) is given by solving the equation $Y(z) = F(m + \sqrt{\alpha r}\, z + \Gamma Y(z))$ with $\Gamma = \gamma - \alpha = \alpha U/(1-U)$. We refer to the term ΓY in the local field as the output proportional term. The appearance of the term in the local field is characteristic of the deterministic analog neural networks, where the TAP equation (1) is , by definition, free of the so-called Onsager reaction field (ORF) which appears in the TAP equation of the stochastic Ising spin neural networks. Note that the Ising spin network with inverse temperature β differs from the analog network with the transfer function $\tanh \beta u$ only by the presence of the ORF-term in the corresponding TAP equation , the equilibrium version of which reads[6]

$$S_i = \tanh \beta\left(\sum_{j=1}^{N} J_{ij}S_j - \frac{\alpha S_i}{1-\beta(1-q)}\right), \qquad q = \frac{1}{N}\sum_{j=1}^{N} S_j^2 \ . \tag{8}$$

with S_i representing thermal average of spin i . Applying the SCSNA to (8) , one can easily recover the well known result of the AGS theory [1] which is obtained by setting $\Gamma = 0$ in (7) [11]. This is because the above mentioned ΓY term is canceled out by the ORF and hence no output proportional term ensues in the local fields of the Ising spin neurons.

The storage capacity of the analog network with analog gain β is a little larger than that of the corresponding Ising spin network with the same β as shown in Fig.1 due to the ΓY term. Another effect of the term is to make the local field distributions of analog networks non gaussian [11,12,13,15].

3. Enhancement of the storage capacity and errorless memory retrieval in networks of nonmonotonic neurons

We apply the SCSNA to investigate the retrieval properties of networks with nonmonotonic transfer functions. One of a typical type of nonmonotonic transfer functions is shown in Fig.2, which is obtained by cutting off the output activity for large |u| of a two-state neuron ($F = \mathrm{sgn}\, u$):

$$F(u) = \mathrm{sgn}\, u \ (|u|<\theta_1), \ \ \frac{-1}{\theta_2-\theta_1}(u - \theta_2 \mathrm{sgn}\, u) \ \ (\theta_1<|u|<\theta_2), \ \ 0 \ (|u|>\theta_2) \quad . \tag{9}$$

The retrieval state will be given by the solution of (7) with $m \neq 0$ which makes the corresponding fixed point of the dynamics (1) stable . We show in Fig.3 the dependence of the order-parameter overlap m satisfying $g \approx 1$ on the storage ratio α (thick line) which was obtained by numerically solving (7) for $\theta_1=\theta_2= \theta = 0.8$. It is noted that the standard type of retrieval solution with $r \neq 0$, in general, is allowed to exist only for a certain interval of α unlike the common case of the Hopfield model with sigmoidal transfer functions. Its upper and lower bounds are denoted by $\tilde{\alpha}_c$ and α_0 respectively.

When the obtained retrieval solutions turn out stable attractors of (1) for α upto $\alpha = \tilde{\alpha}_c$, the $\tilde{\alpha}_c$ coincides with the storage capacity of the network. However, the results of numerical simulations with N = 200 - 800 for various values of α , which are also displayed in Fig.3 , show that whereas the result of the SCSNA is seen to be valid in the case of successful retrieval, the theoretically obtained retrieval solutions with α ranging from $\tilde{\alpha}_c$ down to a certain value α_c, ≈ 0.42

in this case, lose their stability. The occurrence of instability includes that of oscillatory instability, since the present network has no Lyapunov function. The storage capacity α_c given by the upper bound of α of the stability limit turns out, in general, to be lower than $\widetilde{\alpha_c}$ ($\alpha_c \leq \widetilde{\alpha_c}$). Figure 4 displays the phase diagram on the θ - α plane showing the dependence of α_c, $\widetilde{\alpha_c}$, and α_0 on θ in the case of $\theta_1=\theta_2= \theta$ [11]. Enhancement of the storage capacity relative to the commonly known value ≈ 0.14 of the Hopfield model with sigmoidal transfer functions is seen to occur as θ is decreased from $\theta = \infty$.

We present in Fig. 5 the time course of the tolerance overlap g in the retrieval process obtained by numerical simulations with $N = 500$ for $\theta = 0.3$ and $\alpha = 0.3$ ($\alpha < \alpha_0$). It can be seen that the network with the given parameters , if started with an initial condition ensuring an appropriately large overlap, settles into the retrieval state specified with $g = 1$ after some time. This implies that the present system with $\alpha < \alpha_0$ works properly as an associative memory .

The lower bound α_0 in Figs.3 and 4 comes into existence in such a way that with the storage ratio α approaching α_0 from above, r tends to 0, and below α_0 no standard type of retrieval solution can exist. Note that $\sigma^2 = \alpha r$ represents variance of the noise in the local fields.

When we note that $r \to 0+$ implies $|U| \to \infty$ and hence $\Gamma \to -\alpha$, it turns out that $m \to \theta + \alpha/2$ follows from the order parameter equations (7). The straight line $m = \theta + \alpha/2$ drawn (thin line) in Fig.3 shows that the curve representing the standard type of retrieval solution m of (7) just disappears with $r \to 0+$ upon crossing it at $\alpha = \alpha_0$, and that $m(\alpha)$ of successful retrieval with $\alpha \leq \alpha_0$ satisfies the relation $m = \theta + \alpha/2$. This implies that when $\alpha \leq \alpha_0$, the vanishing of noise in the local field of neurons occurs with $r = 0+$ and $\Gamma = -\alpha$. Then we may say that the phenomenon is a result of condensation of the naive noise in the local fields (naive noise $\to$ $-\alpha x_i$) [11,12]. Since one can easily see that $g = 1$ holds exactly when $r = 0+$, the retrieval state with $\alpha \leq \alpha_0$ turns out to ensure memory retrieval without errors. We here note that even if $\theta_2 \neq \theta_1$, the retrieval state with $r = 0+$ can occur for $\theta_2-\theta_1 \leq \alpha \leq \alpha_0$, as long as $\theta_2-\theta_1$ is small. This is because the renormalized output function Y(z) used in (7) for $\theta_1=\theta_2= \theta$ still remains valid within the framework of the SCSNA in the case where $(\theta_2+\theta_1)/2 = \theta$, as long as $\theta_2-\theta_1 \leq |\Gamma|$ [11] . When $\alpha < \theta_2-\theta_1$, the errorless retrieval state is destroyed and a retrieval state with $r \neq 0$ occurs. The details will be studied elsewhere [14,15].

Finally we remark that the enhancement of the storage capacity together with the onset of the $r = 0+$ state, which leads to an improvement of network performances of associative memory are generic [11,12,13,14,15,16,17] and occur for nonmonotonic transfer functions with sufficiently steep negative slopes.

References

[1] D.J. Amit, H. Gutfreund, and H. Sompolinsky, Phys. Rev. Lett. 55, 1530 (1985); Ann. Phys. 173, 30 (1987).

[2] D.J. Amit, Modelling Brain Function (Cambridge: Cambridge University Press) (1989).

[3] J.J. Hopfield, Proc. Natl. Acad. Sci. USA 81, 3088 (1984).

[4] M. Shiino and T. Fukai, J. Phys. A: Math. Gen. 23, L1009 (1990).

[5] R. Kuhn, S. Bos, and J.L.van Hemmen, Phys. Rev. A 43, 2084 (1991).

[6] M. Mezard, G. Parisi, and M.A. Virasoro, Spin Glass Theory and Beyond (Singapore: World Scientific) (1987).

[7] T. Fukai and M. Shiino, Phys. Rev. A 42, 7459 (1990).

[8] F.R. Waugh, C.M. Marcus, and R.M. Westervelt, Phys. Rev. A 43, 3131 (1991)

[9] T. Fukai and M. Shiino , J. Phys. A: Math. Gen. 25, 4799 (1992).

[10] M. Shiino and T. Fukai, J. Phys. A: Math. Gen. 25, L375 (1992).

[11] M. Shiino and T. Fukai, Phys. Rev. E (1993) in press.

[12] M. Shiino and T. Fukai, J. Phys. A: Math. Gen. (1993) in press.

[13] T. Fukai and M. Shiino, preprint.

[14] M. Shiino and T. Fukai, preprint

[15] M. Shiino and T. Fukai, preprint

[16] M. Morita, S. Yoshizawa, and K. Nakano, Trans. IEICE,J73-D-II(2),232 (1990).

[17] S. Yoshizawa, M. Morita, and S. Amari, METR92-01 Technical Report, (1992).

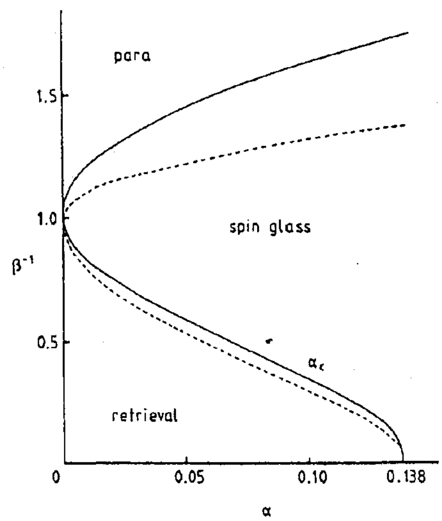

Fig.1 Phase diagram displaying the storage capacities α_c for the analog neural networks with the transfer function tanh βu (solid line) and the Ising spin network with inverse temperature β (dashed line) [4].

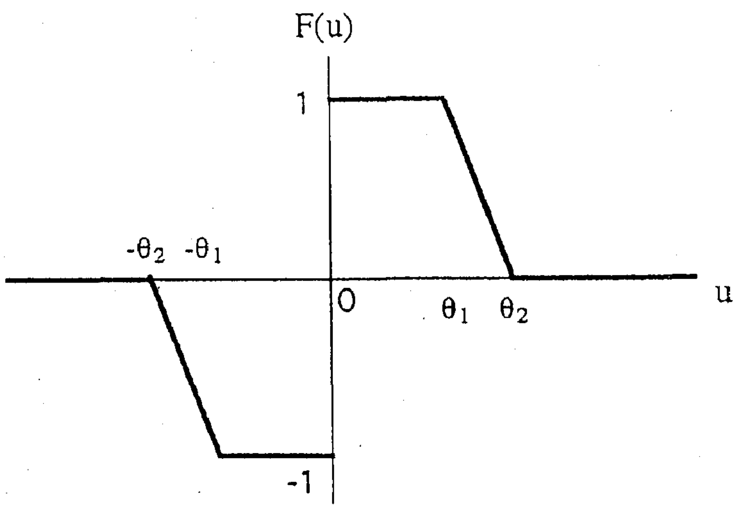

Fig.2 Nonmonotonic transfer function with two parameters θ_1, θ_2..

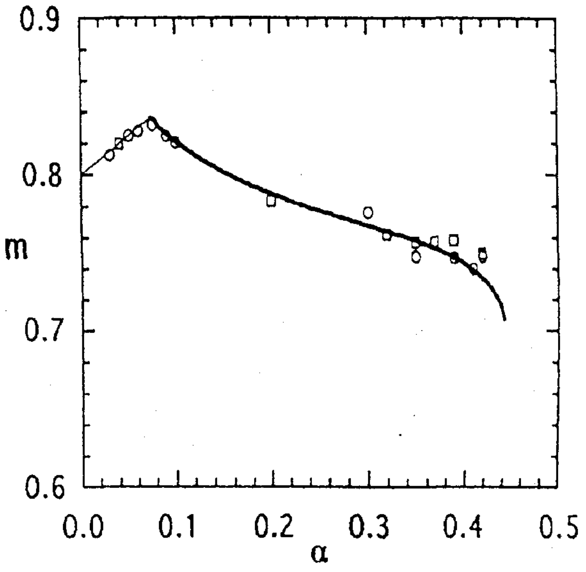

Fig.3 Dependence of m on α for $\theta_1=\theta_2=\theta=0.8$. The thin straight line with $0 \le \alpha \le \alpha_0 (=0.073)$ represents the r = 0+ retrieval phase

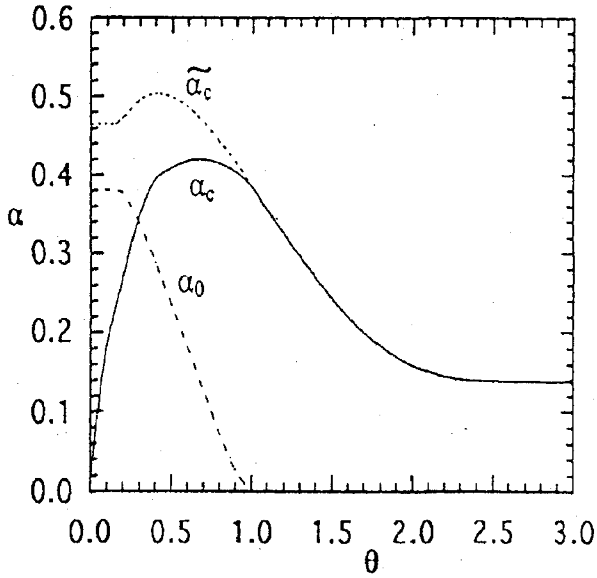

Fig.4 Equilibrium phase diagram in the case of $\theta_1=\theta_2=\theta$.

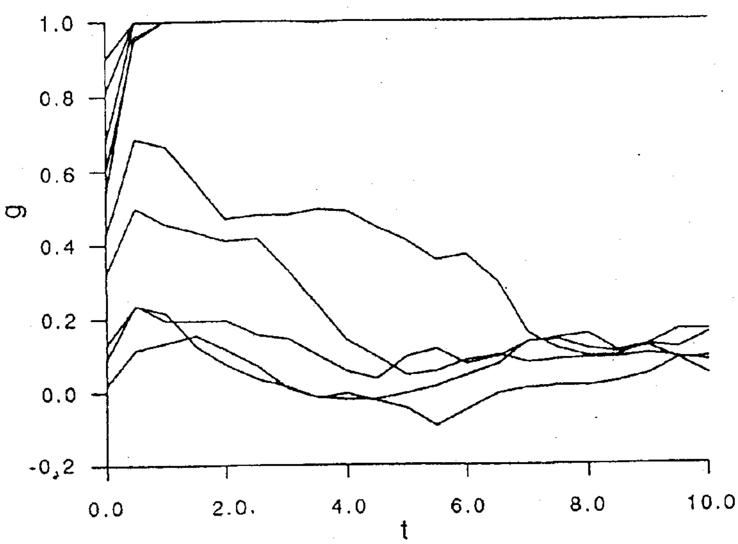

Fig.5 Time course of the retrieval process in terms of g. The equilibrium retrieval state is in the r = 0+ phase with $\theta = 0.3$, $\alpha = 0.3$ (N = 500).